RESEARCH **Open Access**

Seamless clustering multi-hop routing protocol based on improved artificial bee colony algorithm

Tianyi Zhang[1], Geng Chen[1*], Qingtian Zeng[1*], Ge Song[1*], Chao Li[1] and Hua Duan[2]

Abstract

An important issue in the mobile sink wireless sensor networks (MSWSNs) is sensor energy optimization. In order to alleviate the problem of unbalanced network load and high energy consumption in MSWSNs, we proposed a new data collection protocol in this paper: seamless clustering multi-hop routing protocol based on improved artificial bee colony algorithm (IABCP). Because of limited by the communication sensing range and intelligence of ordinary nodes, routing paths can only be constructed by crude methods. And the movement of the sink node will generate a large amount of energy consumption for locating the sink node. In order to solve this problem, we assign the task of routing table generation to the sink node, which will generate the routing table through the improved artificial bee colony algorithm. In addition, we adopt a new method to select cluster head (CH) nodes; node uses the average energy of the surrounding nodes and its own residual energy to calculate the claimed cluster head time. Moreover, we added a sub cluster head CH-β node. When the CH node reaches the number of replacement rounds, the CH-β node becomes CH directly. The simulation results show that our routing protocol is more robust compared with three other protocols.

Keywords: Artificial bee colony algorithms, Load balancing, Network life cycle, Routing protocol

1 Introduction

Wireless sensor networks (WSNs) consist of hundreds or thousands of sensor nodes, it usually includes multiple ordinary nodes and one or more sink nodes. Ordinary nodes are usually arranged in space by the random or artificial way, and cannot move after being arranged. Generally, ordinary nodes are powered by their own batteries, which cannot be replenished, and periodically or event triggered to collect the information of the surrounding environment and transmit it to the sink node. The sink node can move or stay freely in a wireless sensor network. Compared with ordinary nodes, the energy of the sink node can be regarded as infinite, and the sink node usually has higher computing power. Therefore, the enhancement of lifetime and energy efficiency is an important issue.

In recent years, many researchers have proposed various routing protocols of WSNs to alleviate network load imbalance and improve network lifetime. Network load unbalance is a widespread problem in WSNs, and it is also a problem that we want to avoid as much as possible. The uneven energy consumption of each round of nodes is the cause of network load imbalance. In WSNs, there are two main routing methods: single-hop routing method and multi-hop routing method. In single-hop routing and multi-hop routing, the load of ordinary nodes is often related to their location. Single-hop clustering routing such as low-energy adaptive cluster hierarchical (LEACH) [1], energy efficient clustering scheme (EECS) [2], and most of the improved protocols based on LEACH [3] have the problem of higher node load far from the sink node. In multi-hop networks, such as mobile sink-based data gathering (MSDG) [4] and economic-environmental unit-commitment (EEUC) [5] routing protocol, it is inevitable that nodes in a hop range of communication radius of sink

* Correspondence: gengchen@sdust.edu.cn; qtzeng@sdust.edu.cn;
songge@sdust.edu.cn
[1]College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao 266590, China
Full list of author information is available at the end of the article

nodes are required to act as relay nodes in routing paths, and nodes close to sink nodes have higher load. If the sink node moves continuously and traverses the whole network for a long enough time, the relative position of each node to the sink node will change constantly, and the unbalanced state of load in the sensor network will be alleviated. Therefore, MSWSNs [6] was born.

As mentioned above, although the mobility of the sink node alleviates the network load imbalance [7, 8], it also brings some problems. For example, in two-tier data dissemination (TTDD) protocol [9], the mobility of sink nodes results in the failure and extension of routing path. The failure of the routing path leads to the loss of data. And the prolonged routing path will lead to more hops and higher network energy consumption. Routing failure is mainly caused by the poor update timeliness of the routing table. The location of the sink node has changed, but the rest of the nodes still send messages according to the original routing table, so the data cannot reach the sink node and lose. At the same time, it also consumes a lot of energy in MSWSNs to synchronize the location of sink nodes.

In this paper, in order to solve the problem of high energy consumption and load unbalance of MSWSNs, we propose a new clustering routing protocol for large-scale MSWSNs. Limited by the communication sensing range and intelligence of ordinary nodes, routing paths can only be constructed by crude methods in traditional routing protocols [10–12]. Therefore, in the construction of the routing path, the balance of the network load cannot be considered. And with the moving of the sink node, it will consume a lot of energy for each node to synchronize the location of the sink node. In our protocol, we use the sink node to generate routes, which will generate the routing table through the improved artificial bee colony algorithm (IABC) [13]. The generated routing table considers both the overall energy consumption of the network and the balance of the network load. And in the whole protocol flow, the sink node only needs to know the location of the first-round cluster head. In addition, we adopt a new method to select CH nodes. Each node uses the average energy of the surrounding nodes and its own residual energy to calculate the claimed cluster head time. Moreover, a sub cluster head CH-β node is added between the cluster head and the member node. At the beginning of the network, the CH node is selected first and then the CH-β node. After CH runs to the set number of rounds, the CH-β node directly replaces CH and becomes a new CH. Then the network selects CH-β nodes according to the CH election method. In the cluster, CH and CH-β nodes exist at the same time. The cluster head can add the position of CH-β node

into the data packet and pass it to the sink node so that the sink node can obtain the position information of the next round of CH. In this way, the node can get the corresponding routing path generated by the sink node without knowing the location of the sink node. Compared with LECAH protocol, MSDG protocol, EEUC protocol, and the original bee colony algorithm, the protocol alleviates the load imbalance of the network and prolongs the service life of the network. Moreover, the optimization of the bee colony algorithm improves the efficiency of protocol routing construction.

2 Related work

LEACH protocol is the earliest clustering protocol proposed. It proposes a method to reduce network energy consumption by clustering. The protocol uses the method of generating random numbers to declare cluster heads (CH). Ordinary nodes join the nearest CH as sub-nodes. The cluster head processes the data of the nodes in the cluster and forwards the data to the sink node. Clustering reduces the overall energy consumption of sensor networks [14], but clustering also causes other problems. LEACH protocol can only adjust the size and number of clusters in the network by adjusting the probability of nodes declaring CH. Although nodes cannot continuously act as CH, the geographical distribution of CH is uncertain, which makes the lifetime of LEACH protocol unstable [15]. Moreover, the LEACH protocol communicates with sink nodes in a single-hop mode, which has poor adaptability to the network environment and short lifetime in large networks. This is also a problem that LEACH protocol and most improved protocols based on LEACH protocol [16] cannot overcome.

Hybrid energy efficient distributed (HEED) [17] protocol begin considers the influence of the residual energy of the nodes and the maximum network energy of the WSNs. Therefore, a node with higher residual energy has more chances of being selected as the CH which improves the stability of the network and reduces the energy consumption of the network. Therefore, a node needs to obtain the energy information of all other nodes to determine its clustering radius. In the cluster head election process, each node needs a full-network broadcast. When the range of the wireless sensor network is large, the network energy consumption is high.

Rotated unequal hybrid energy efficient distributed (RUHEED) [18] protocol is an improvement of HEED, in order to solve the problem of high energy consumption in the large-scale network, RUHEED protocol introduces the concept of "rotation." After the tenure of the cluster head node arrives, the cluster head node specifies

the node with the largest energy as the successor node without any judgment. This does avoid the process of cluster head election and reduce the energy consumption in the cluster head stage. However, when the nodes die, the network will need to re cluster. When the energy of nodes is generally small, the election of CH will inevitably accelerate the death of nodes. Even after the death of a single node, the whole network will fall into the process of campaign cluster head-death-campaign cluster head-death.
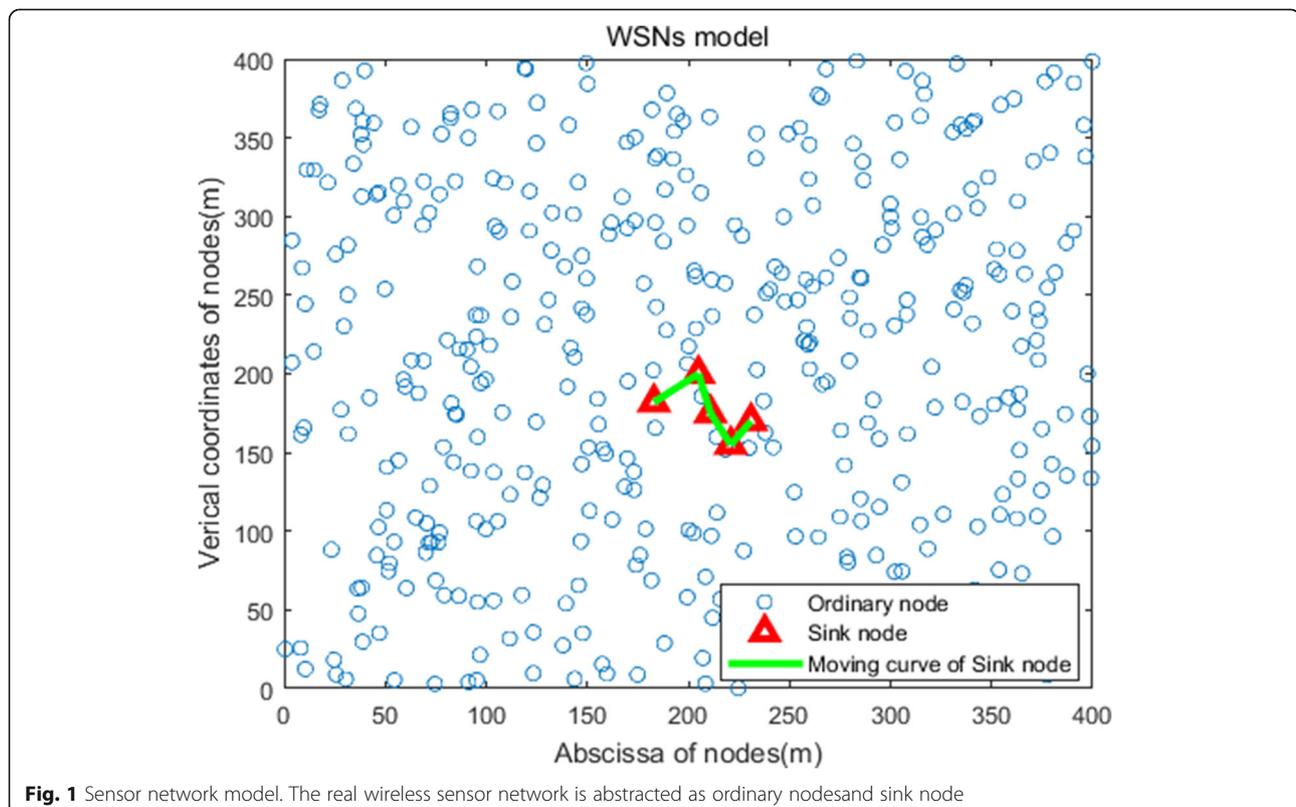
EECS protocol proposes a time broadcast method to make the network form a cluster with approximately uniform distribution, so that the selection of CH can get rid of randomness. Ordinary node chooses the cluster head by the distance between cluster head nodes and sink node, so that cluster head near the sink node has more sub-nodes. EECS adopts a non-uniform clustering method to balance the network load.
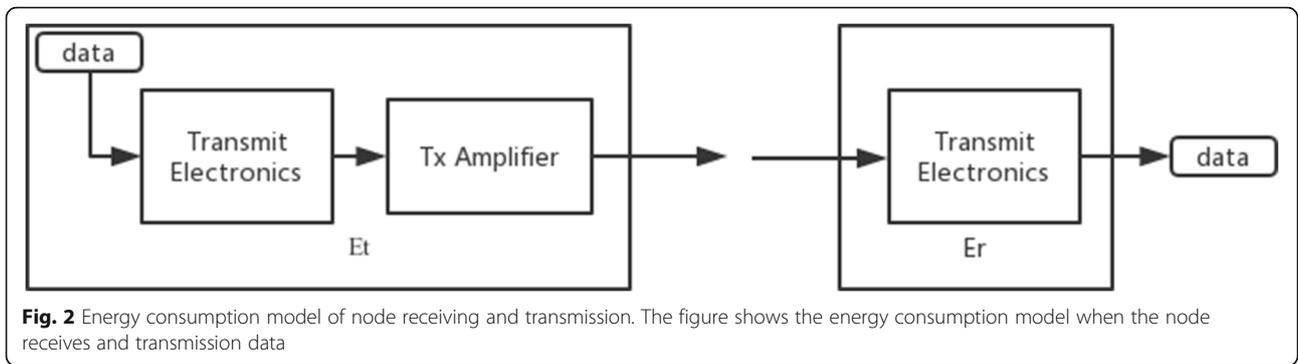
Power efficient gathering in sensor information system (PEGSIS) [19] protocol is the predecessor of multi-hop routing. It reduces the overall energy consumption of the network by linking nodes into chains and choosing individual nodes to communicate with sink nodes, but it also brings about the problem of high network delay.

Therefore, a real multi-hop routing protocol, such as MSDG and EEUC protocol, is generated. MSDG protocol transmits data to the sink node by clustering and tree formation. However, in the face of a high-density network environment, its load balance is poor, and it is easy to cause network failure due to energy holes. Although EEUC balances the network load through non-uniform clustering, its clustering and routing path construction process consume more energy, and cannot guarantee the load balance of the constructed routing, so its lifetime is also affected.

Mobile sink-based routing protocol (MSRP) [20] adopts the method that the mobile sink node close to each CH to collect data. Although this approach makes the energy consumption of the whole network relatively low. But obviously, this method is not suitable for real-time applications of WSNs.

In recent years, many new routing protocols have begun to apply mature path search algorithms to the routing construction of WSNs, such as Adaptive Periodic Threshold-sensitive energy-efficient sensor network (APTEEN) [21] protocol uses the ant colony algorithm to build multi-hop routing. And, PEGASIS in 3DWSN based on genetic algorithm (PEG-GA) [22] protocol improves the PEGASIS protocol in the 3D application

Fig. 1 Sensor network model. The real wireless sensor network is abstracted as ordinary nodesand sink node

Fig. 2 Energy consumption model of node receiving and transmission. The figure shows the energy consumption model when the node receives and transmission data

environment, using a genetic algorithm instead of the greedy algorithm to link nodes into the chain to generate the node chain with the minimum length. But they all require ordinary nodes to have strong computing power. This will increase the cost of the network.
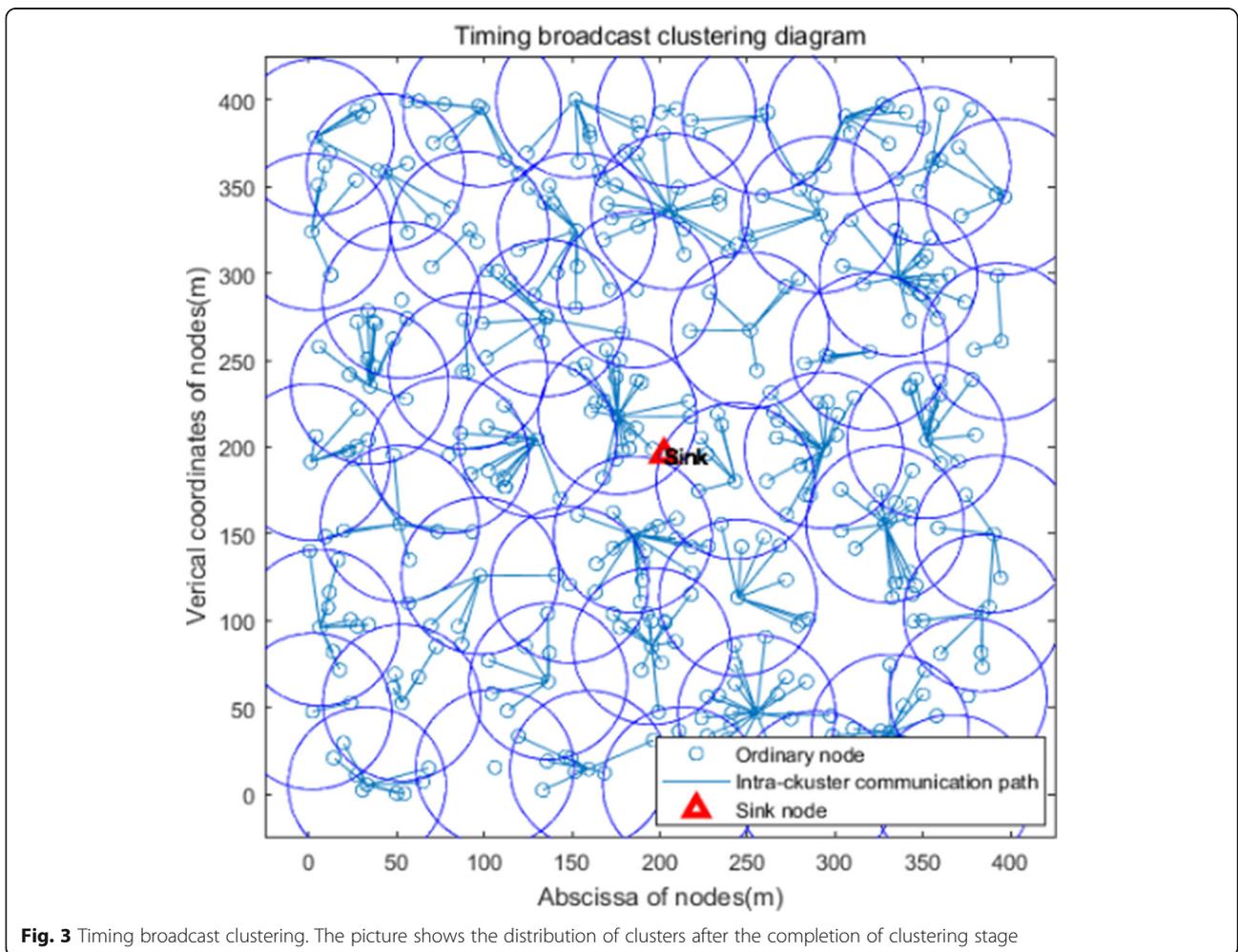
The rest of this paper is organized as follows. Section 3 presents the system model. Section 4 proposed the protocol flow, including the seamless clustering process and routing construction process. Section 5 presents the simulation results. Section 6 draws the conclusions.

3 System model

3.1 Sensor network model

The sensor network model is as follows in Fig. 1.

Fig. 3 Timing broadcast clustering. The picture shows the distribution of clusters after the completion of clustering stage

(1) The network consists of a sink node and n ordinary nodes.

(2) Ordinary nodes are randomly distributed in the region, and the initial energy is isomorphic and cannot be supplemented.

(3) Ordinary nodes are no longer moved after they are laid out, and nodes can get their positions through GPS.

(4) The communication power of ordinary nodes can be adjusted according to the communication distance.

(5) Cluster head nodes use a data fusion method to process the data of nodes in the cluster.

(6) The initial position of the sink node is located in the sensor network and can move randomly at a constant speed in the sensor network layout area.

(7) Mobile sink node energy can be replenished at any time.

(8) Nodes periodically perform data acquisition and always send data to sink nodes.

(9) Each node has its own unique ID.

3.2 Node energy consumption model

Node energy consumption includes two parts: data receiving energy consumption and data sending energy consumption. Data receiving energy consumption mainly includes receiving circuit energy consumption. Data sending energy consumption includes transmitting circuit energy consumption and power amplifier circuit energy consumption. The energy consumption of the node receiving and transmission model is shown in Fig. 2. The energy consumption of node receiving and transmission data is shown in Eqs. (1) (2).

$$E_r = \text{data} \times E_{\text{elec}} \tag{1}$$

$$E_t = \begin{cases} \text{data} \times \left(E_{\text{elec}} + \varepsilon_{\text{fs}} \times d^2\right), d \le d_0 \\ \text{data} \times \left(E_{\text{elec}} + \varepsilon_{\text{mp}} \times d^4\right), d > d_0 \end{cases} \tag{2}$$

ε_{fs} and ε_{mp} are the energy consumption coefficients of power amplifier circuits under different channel propagation models, d denotes the distance of data single-hop transmission, and d_0 is the switching threshold of amplifier circuits determined by Eq. (3). In order to simplify the process, data is divided into control message (CM) and data message (DM). The DM packet is the data information generated by the ordinary node. The other short packet data communication used for routing and ACK code are regarded as CM packets.
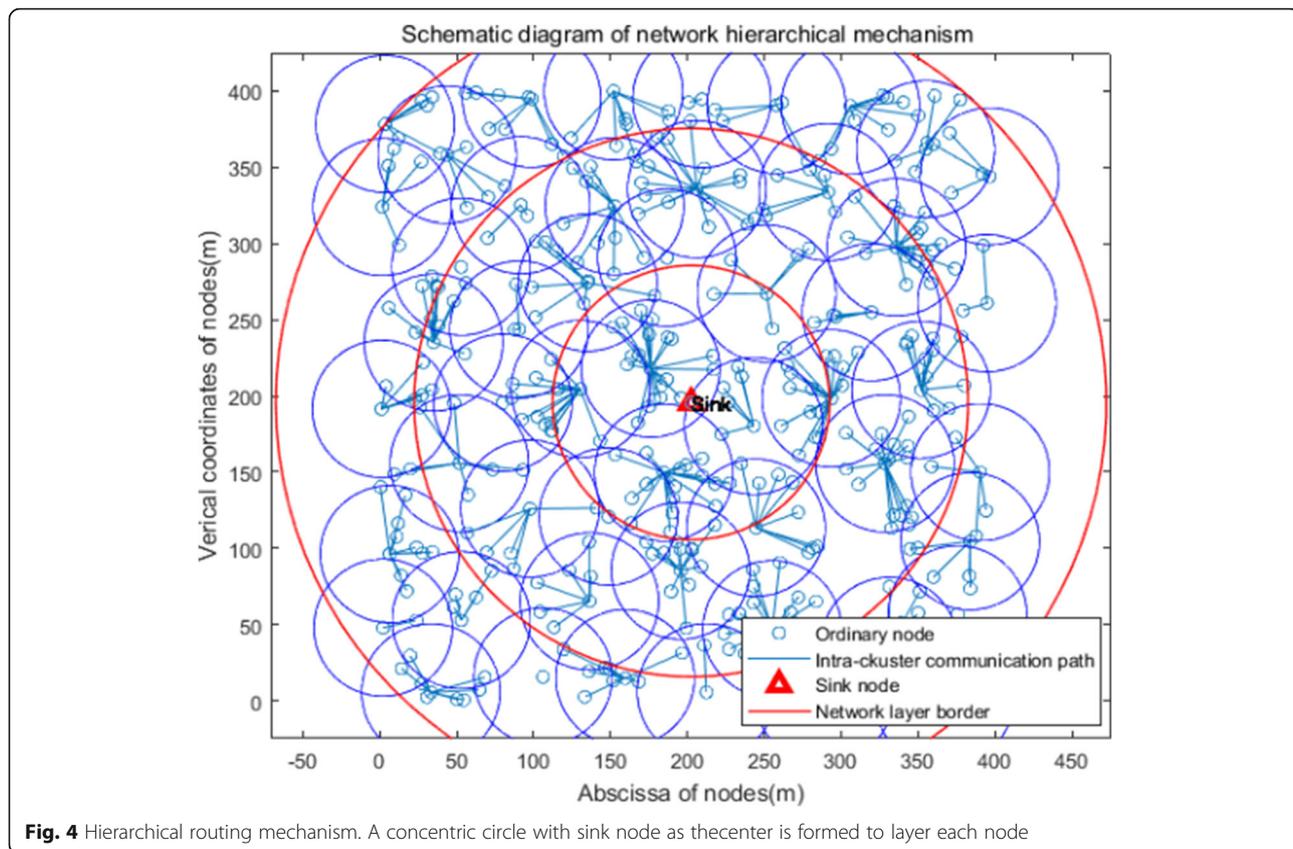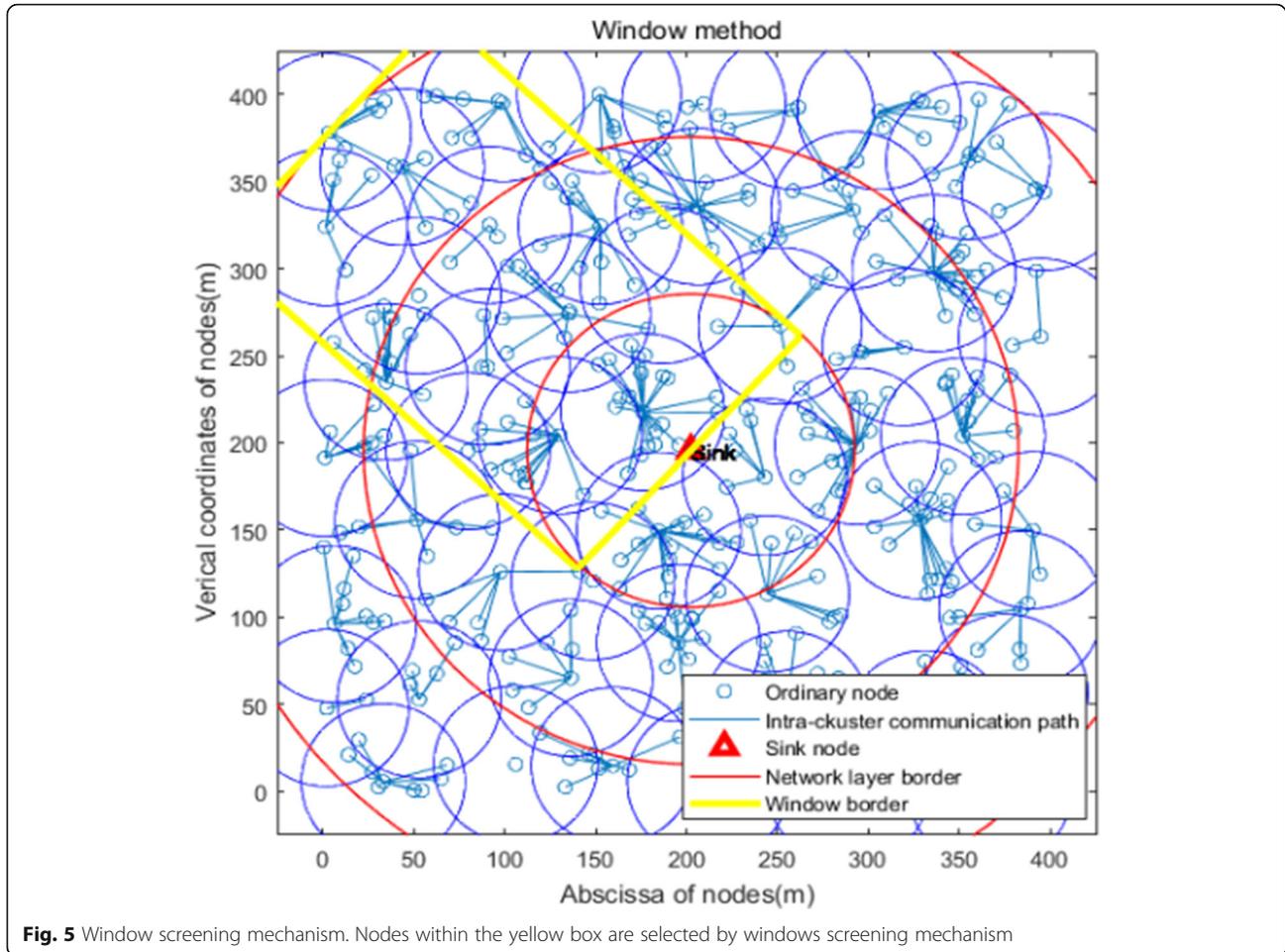

Fig. 4 Hierarchical routing mechanism. A concentric circle with sink node as the center is formed to layer each node

Fig. 5 Window screening mechanism. Nodes within the yellow box are selected by windows screening mechanism

$$d_0 = \sqrt{\frac{\varepsilon_{\text{fs}}}{\varepsilon_{\text{mp}}}} \tag{3}$$

Moreover, cluster head nodes consume additional energy for data fusion. Cluster head nodes compress and pack the collected information of cluster nodes, which reduces the total amount of data flowing into the network. The process of packaging and compression is called data fusion. In order to simplify the design, this paper assumes that the cluster head will compress all DM emitted by the collected intra-cluster nodes into the DM size, and the energy consumption of data fusion is set to E_{DA}= 5 nJ/bit. The energy consumption of data fusion$E_f$is shown in Eq. (4).

$$E_{\text{f}} = E_{\text{DA}} \times \text{data} \tag{4}$$

4 Protocol design

In MSWSN, routing path failure and routing path extension are often caused by the movement of sink nodes, which leads to data loss and network energy consumption increased. In addition, ordinary nodes also need to update the location of sink nodes in real time. It often uses the way of sending messages from the sink node and forwarding messages to the whole network through the surrounding nodes step by step, which will bring great energy consumption. In order to resolve this problem, we design a new protocol: IABCP. In our protocol, we give the routing path generation task to the sink node, which generates the routing path through the improved bee colony algorithm. Therefore, the ordinary node does not need to synchronize the location of the sink node. And in the whole protocol flow, the sink node only needs to know the location of the first-round

Table 1 Network environment simulation prerequisites

Parameter	Value	Parameter	Value
Network coverage	$400 \times 400 m^2$	d_0	90
Node number	400–2000	CM size	200 bit
Initial energy	0.25 J	ε_{fs}	10p J/bit/m^2
DM size	4000 bit	ε_{mp}	0.0013 pJ/bit/m^4
E_{elec}	50n J/bit	E_{DA}	5 nJ/bit

cluster head to generate a routing table starting from itself. This is also the reason why the node needs to be equipped with GPS. Compared with the ordinary node, the sink node has a broader vision and stronger computing power. It can generate routing paths while taking into account the overall energy consumption and load balance of the network. However, the cluster head will change after running to the set number of rounds. In order to avoid the process of new cluster heads informing the sink node of its position when cluster head changes, we add sub cluster head CH-β node between cluster head node and member node. CH node and CH-β node are selected in order at the beginning of the network. At the beginning of the network, the CH node is selected first and then the CH-β node. After CH runs to the set number of rounds, the CH-β node directly replaces CH and becomes a new CH, and then the network carries out the campaign of CH-β node. Therefore, CH-β node and cluster head node exist at the same time in the subsequent inter cluster routing. The cluster head node can include the position of CH-β node in the packets and transferred to the sink node. The sink node can calculate the time of cluster head change and calculate the new routing table according to the location of the new cluster head. And because the CH-β node becomes the cluster head directly when the CH change, the seamless clustering is realized in the network. It will reduce the empty window period of network data collection. In the overall process of protocol, the behaviors of

sink nodes and ordinary nodes are shown in Algorithm 1 and 2:

Algorithm 1: Ordinary node behavior	**Algorithm 2**: Sink node behavior
Input: Running times, Routing table message	**Input:** Running times, Sink node location
Output: Void	**Output:** New routing table
1 **if** Running times ==1	1 **if** The distance from the last generated routing table exceed the cluster radius
2 Node election CH and CH-β node;	2 r=1;
3 CH determines its own number of turns through network hierarchical mechanism;	3 **end if**
4 CH sends routing table request information to sink node through directed flooding;	4 **if** Running times == CH change time
5 **end if**	5 r=1;
6 **if** Network is still operational	6 **end if**
7 **if** CH hasn't reached change time	7 **if** Sink node receive routing table request information from the CH;
8 **if** CH receives new routing table	8 r=1;
9 CH update routing table;	9 **end if**
10 CH sends data according to routing table;	10 **if** r==1
11 **else**	11 Sink node reproduce the routing table by IABC algorithm and forward it back by routing table;
12 CH sends data according to routing table;	12 **end if**
13 **end if**	
14 **else**	
15 CH-β node upgrade to CH;	
16 Node election CH-β node;	
17 **end if**	
18 **end if**	

When the sink node generates a new routing table, it flips the routing table. The original routing table that starts with the sender node and ends with the sink node is transformed into a routing table that starts with the sink node and ends with the sender node. The sink node sends the information of the original routing table to each node according to the routing table.

Clustering method will be introduced in detail in section 4.1. The sink node generates routing tables through an improved bee colony algorithm. The bee colony algorithm and its improved method will be introduced in detail in section 4.2.
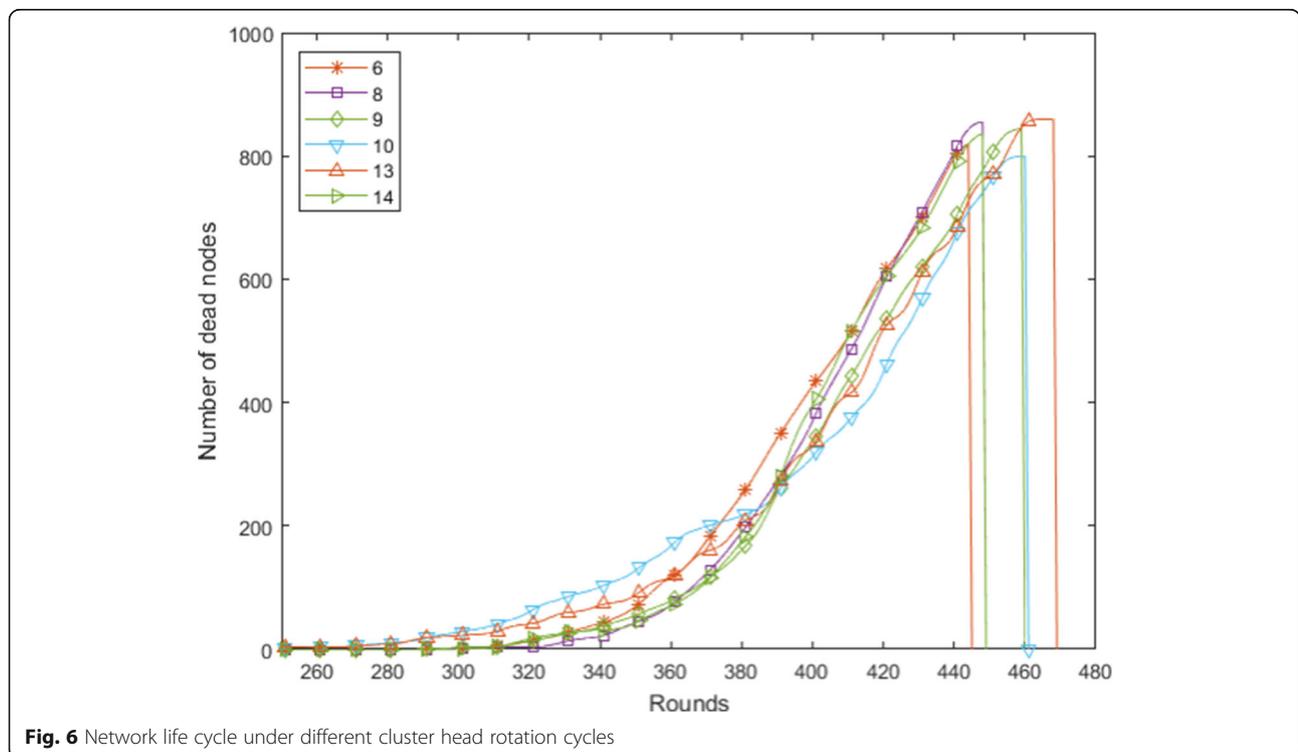

Fig. 6 Network life cycle under different cluster head rotation cycles

4.1 Seamless clustering of nodes

This protocol divides nodes into cluster heads, ordinary nodes, and CH-β nodes, which are responsible for different tasks. In WSN, the cluster head is responsible for collecting data in cluster and communicating among clusters, the CH-β node is the second node after the cluster head. When the number of rounds run by the cluster head reaches the set value, the CH-β node upgrades directly from the CH-β node to the cluster head. Because there are no cluster heads and CH-β nodes in the initial clustering of networks, it is necessary to elect both cluster heads and CH-β nodes. In the subsequent stage, only CH-β nodes need to be elected in the process of inter-cluster communication. Nodes elect cluster heads by broadcasting the declared information of cluster heads. The time T_i of nodes sending declared information is determined by Eq. (5). Among them $\overline{E}_j$ is the average energy of nodes around node i, E_i is the energy of

nodes themselves, and T_h is the time constant. n is to increase the random disturbance value in order to avoid the situation that multiple nodes simultaneously declare the cluster head. The value of n ranges from − 1 to 1% of the initial energy.

$$T_i = T_h \times \frac{\overline{E}_j + n}{E_i} \tag{5}$$

It is beneficial for the load balance of the whole network to use the nodes with higher energy as cluster heads. Therefore, we make the nodes with higher energy have smaller T_i when running for cluster heads, which makes it easier to become cluster heads. We compare the energy of each node with the average energy of the surrounding nodes to select the nodes with higher energy. Under the premise of avoiding frequent network communication, this method can be used to compare the energy of adjacent nodes.
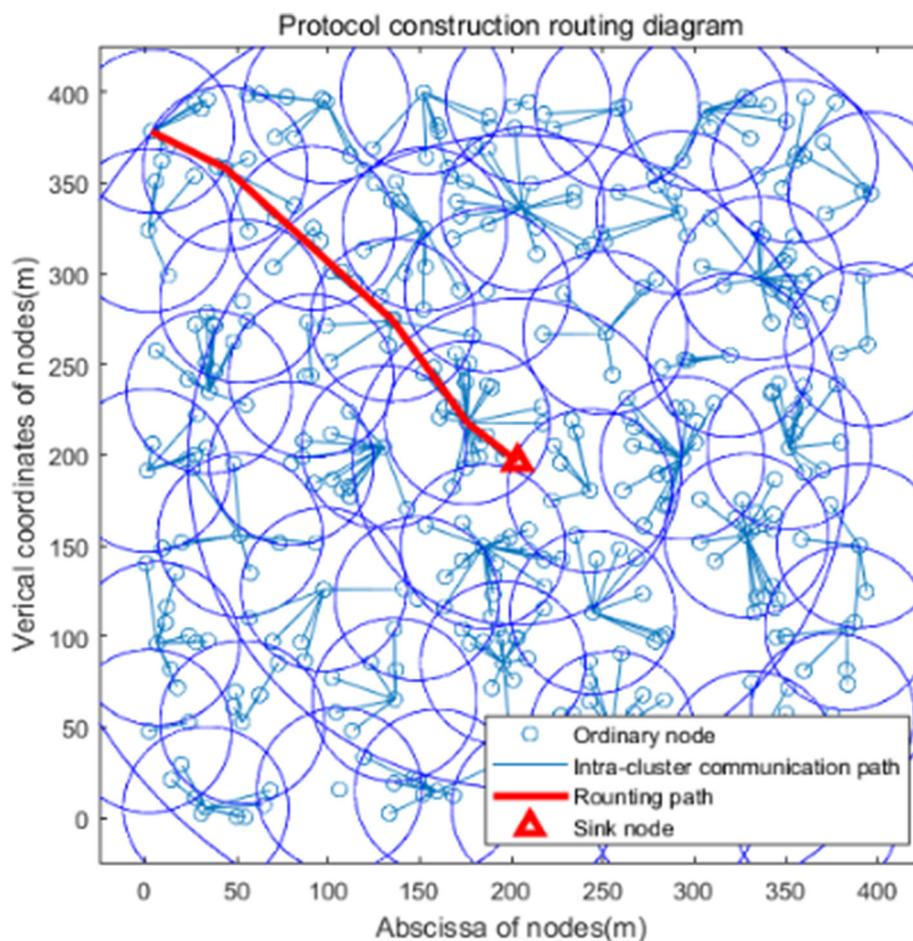

Fig. 7 Protocol construction routing diagram. The red curve is the data transmission route from the sending cluster head to the sink node

The campaign for CH-β node occurs when the cluster head receives the information within the cluster once. At this time, the cluster head waits for the routing table from the sink node and participates in inter-cluster communication. The ordinary node is idle and has enough time to elect CH-β node. The current cluster head does not participate in the campaign for CH-β node, only records the number and position of CH-β node after the election. Because CH needs to bear greater communication pressure and higher energy consumption of nodes, the non-overlap between CH and CH-β nodes makes it impossible for nodes to become CH continuously. The generation of CH-β node makes the replacement of the cluster head known to the sink node. In the process of subsequent routing table construction, the sink node can generate the routing table in advance, and the new cluster head does not need to inform the sink node of its location through broadcast and forwarding, which reduces the complexity and energy consumption of network communication. By upgrading the network from CH-β node to cluster head directly at the end of the cluster head term, the network can be seamlessly clustered. There is no blank area in most existing protocols because of the lack of communication in cluster head election. The concrete steps of the nodes clustering algorithm are pseudo-code as shown in Algorithm 3.

Algorithm 3: Clustering of nodes

Input: Node position, Node energy, Total running rounds of network "N"

Output: Cluster head number

```
1    for i←1 to node number do                          11    else if (j >0) & (k ==0)
2        Node (i) broadcast CM (ID, Energy) to its      12        Node (i) claim to be CH-β node;
         communication range and receive CM from        13    else
         neighboring nodes;                              14        Node (i) becomes a sub-node of CH;
3        Node (i) calculate the average energy of       15    end if
         neighbor nodes and generating random           16    else if (N >1) & (current time= T_i )
         number n;                                       17        k = Number of CH-β claims received by the
4        Node (i) calculate T_i by formula (5);                   Node (i);
5    end for                                             18        if (current time= T_i )&(k==0)
6    if (N ==1) & (current time= T_i )                   19            Node (i) claim to be CH-β node;
7        j = Number of CH claims received by the         20        else
         Node (i);                                       21            Node (i) join declared nodes as sub-nodes
8        k = Number of CH-β claims received by the       22        end if
         Node (i);                                       23    end if
9        if j ==0
10           Node (i) claim to be CH;
```

As shown in Fig. 3, the clustering results show that each cluster is approximately evenly distributed within the network coverage, and it can ensure that any cluster head can maintain communication with at least one cluster head within twice the cluster radius, that is, there is no isolated cluster head in the network, and the communication distance between CH can be indirectly adjusted by adjusting the cluster radius.

4.2 Improved artificial bee colony algorithm

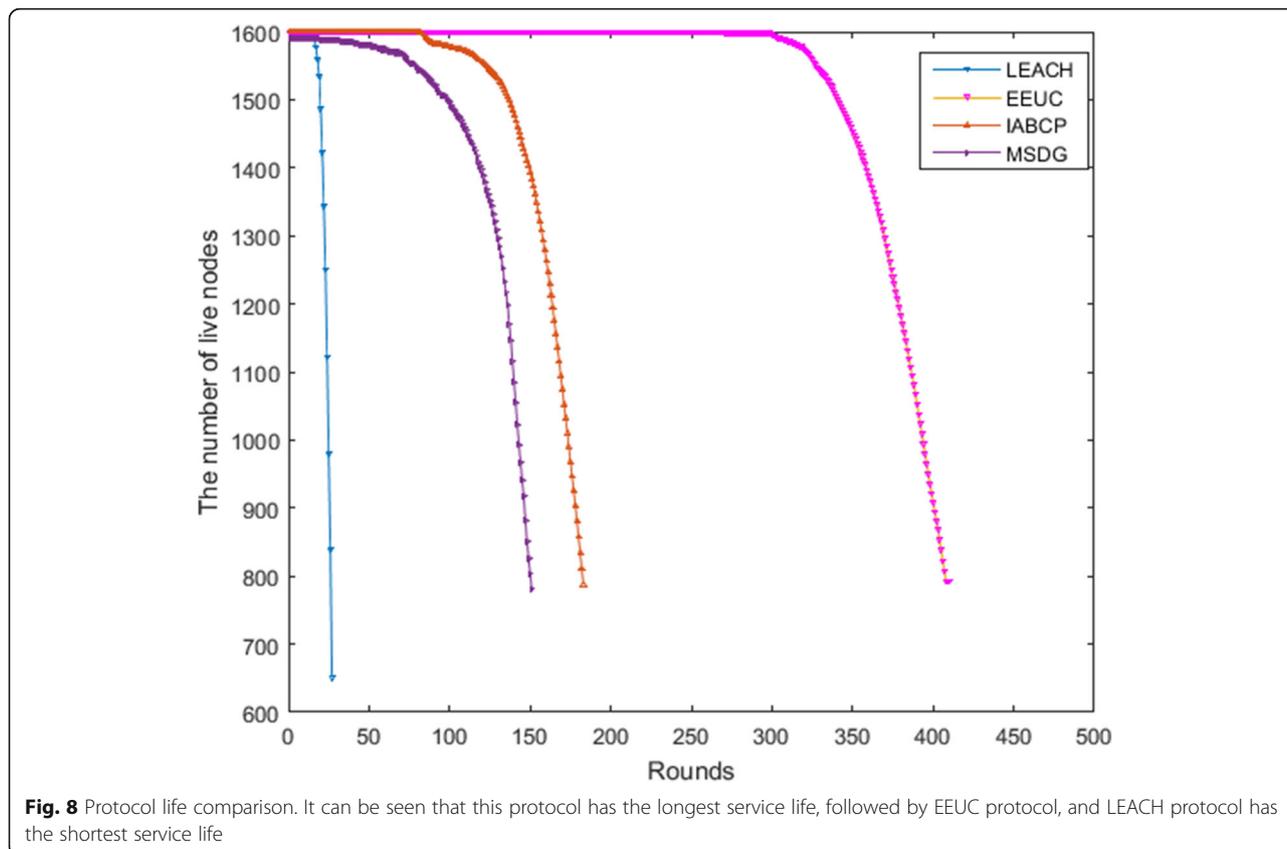In this protocol, we use the artificial bee colony (ABC) algorithm to generate the routing table by the sink node.

Fig. 8 Protocol life comparison. It can be seen that this protocol has the longest service life, followed by EEUC protocol, and LEACH protocol has the shortest service life

ABC algorithm is a kind of swarm intelligence algorithm based on biology. The algorithm seeks the optimal solution by simulating the behavior of obtaining the optimal honey source in the process of collecting honey. In this protocol, ABC algorithm is combined with network hierarchical mechanism and window screening mechanism to improve its initial solution generation and subsequent optimization process, and to enhance network load balancing by improving fitness function. The optimized ABC algorithm flow is shown in Algorithm 4.

Algorithm 4: Improved bee colony algorithm

Input: Cluster Head Node Location

Output: Routing table

```
1    for i←1 to Cluster head node number do
2        for r←1 to The number of rounds set by the algorithm do
3            if node call==0
4                Screening Nodes by window method;
5                Layering nodes through network hierarchical mechanism;
6                Generate initial solution set;
7                The leaders searches perform neighborhood search and greedy selection through improved
                 fitness function;
8                Followers follow leader through fitness selection;
9                Follower performs neighborhood search and greedy selection through improved fitness
                 function;
10               if   The solution is not optimized under the set number of rounds
11                   Random new solution of scouter replaces the original solution;
12               end if
13               Record the solutions corresponding to the highest fitness values generated in each round;
14           end if
15       end for
16       Output the solution with the highest fitness under the set number of rounds as the Routing table
17       for n←1 to Nodes in Routing Table do
18           node call=1;
19       end for
20   end for
```
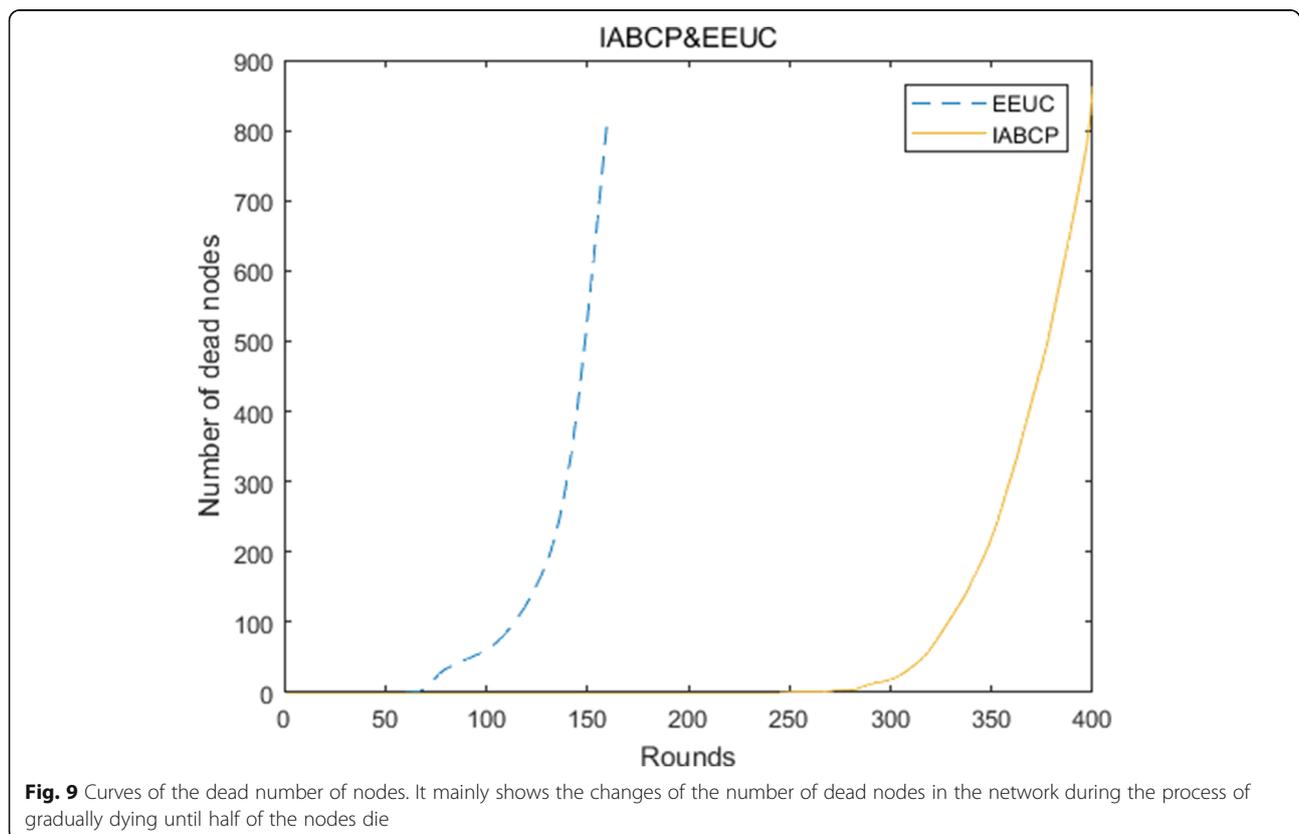

The principle of the ABC algorithm is to optimize the solution; the algorithm constantly generates new solutions through the old ones and replaces the old solution with the new solution with higher fitness value. And the solution with high fitness value has higher optimization probability. In this way, the fitness value of the generated solution becomes higher and higher. Finally, the optimal solution of the corresponding problem can be found in a long enough number of rounds. In order to increase the efficiency of the ABC algorithm, we add window method and network hierarchical mechanism to the original ABC. Network hierarchical mechanism and window method will be introduced in 4.2.1 and 4.2.2 respectively. At the same time, the fitness function is optimized to enhance the load balance of the network. The optimization of the fitness function will be introduced in detail in 4.2.3.

4.2.1 Network hierarchical mechanism

In order to reduce unnecessary energy consumption and improve the efficiency of ABC algorithm, we adopted the network hierarchical mechanism [23, 24] to optimize ABC algorithm. The function of routing layering is mainly reflected in two points:

(1) When the network is clustered for the first time, the sink node sends the control information containing the number of turns. After receiving the

Fig. 9 Curves of the dead number of nodes. It mainly shows the changes of the number of dead nodes in the network during the process of gradually dying until half of the nodes die

Zhang *et al. EURASIP Journal on Wireless Communications and Networking* (2020) 2020:75

Page 11 of 20

information of the number of turns, the CH node compares the number of turns *J* in the packet with the number of turns *K* recorded by itself. If *J > K*, the node discards the packet. If *J < K*, the node updates its own turns information and makes *J = J + 1* continue to forward. In this way, each cluster head knows its own number of turns. When sending its own position to sink nodes in the future, it can guide the direction of data forwarding by limiting the information of high cycles to be forwarded only by nodes with low cycles, avoiding unnecessary flooding energy consumption.

(2) When the sink node builds the route through ABC algorithm, the sink nodes layer each node by the distance between each cluster head and itself. When the initial solution is generated, each node can only transfer data to the layer closer to the sink node. In this way, the generation of invalid solutions is avoided. Moreover, sink node determines the priority of routing table generation by the distance between each node and sink node. The farther the node is from the sink, the higher the priority of the routing table generation. The sink node will no longer generate a routing table starting from the node that the peripheral node

routing table passes through, reduces the total number of generated solutions.

The schematic diagram of the network hierarchy is shown in Fig. 4.

4.2.2 Windows screening mechanism

Network hierarchical mechanism avoids the formation of invalid solutions in ABC algorithm. However, when ABC algorithm searches for solutions, the number of possible solutions is the product of the number of nodes in each layer. Reducing the number of optional nodes in each layer can effectively reduce the number of possible solutions and speed up the efficiency of bee colony algorithm in searching the optimal solution. So, we add the window screening mechanism [25, 26] to ABC algorithm. ABC algorithm can only select the node in the window range when generating the initial solution and optimizing the solution. The schematic diagram is shown in Fig. 5.

Theoretically, the minimum energy consumption and the minimum delay path are a straight line connecting the sink node and the sender node. However, in practice, nodes with a forwarding function are not necessarily distributed on this line, and the location of the distribution
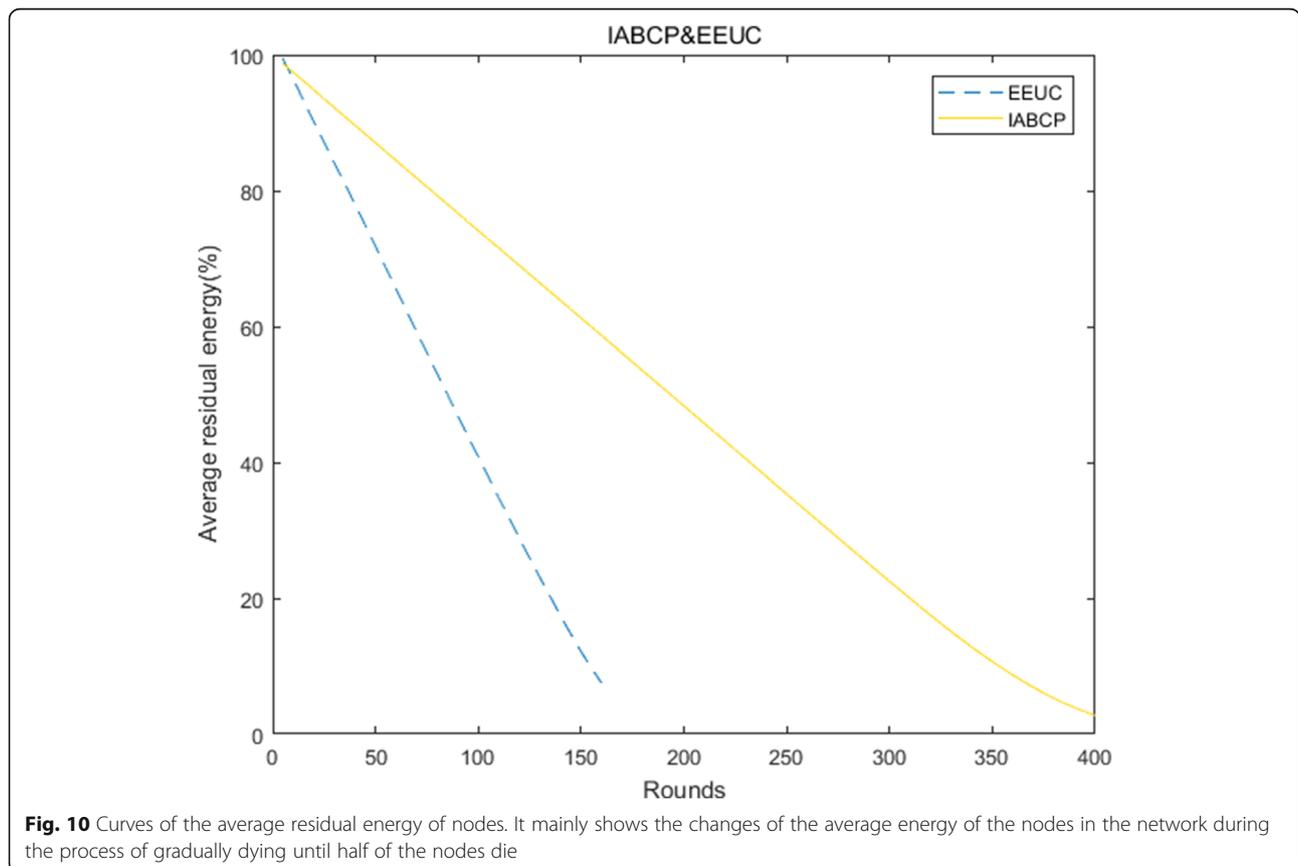

Fig. 10 Curves of the average residual energy of nodes. It mainly shows the changes of the average energy of the nodes in the network during the process of gradually dying until half of the nodes die

does not necessarily meet the minimum energy consumption condition. But this line can be used as the condition of guiding and filtering. As shown in Fig. 5, the central axis of the window is a straight line starting from the sink node and ending at the sender node. According to different requirements, the window size can be adjusted by adjusting the window width. In the case of low communication delay as the main objective, as shown in Fig. 5, the window width is taken as the communication radius of the nodes, and the number of possible solutions in the window is only 45, compared with about 3000 possible solutions before increasing the window method, which is only about 1.5% of the previous ones. The size of the window range has a great influence on the generation and optimization of the solution. When the window range is too small, it is easy to have no alternative solution or local optimal solution. If the window size is too large, the algorithm often needs more rounds to get the optimal solution. Therefore, the size of the window needs to be adjusted according to different optimization requirements. For example, in the following fitness function of this paper, the goal of enhancing network load balancing is added. It is necessary to increase the size of the windows

appropriately. The specific size of windows in different environments needs to be obtained through simulation experiments.

4.2.3 Fitness function optimization

In ABC algorithm, the fitness function is used to evaluate the fitness of the solution, and the solution with low fitness is eliminated to retain the solution with high fitness. The fitness function of the original ABC algorithm is only related to the length of the generated path. However, when constructing the routing path by this method, the task of the algorithm includes not only reducing the network communication delay, but also balancing the network load. Therefore, the fitness function is composed of the average energy of the routing path, the minimum energy of the path, and the similarity between the generated path and the shortest path. The fitness function is shown in Eq. (6).

$$F_i = \frac{\overline{E}_i E_{\min}}{L_i} \tag{6}$$

Among them, $\overline{E}_i$ and $E_{\min}$ are the average energy and the minimum energy of the nodes on the route respectively, and L_i is the length of the route. That is to say,

Fig. 11 Variance of residual energy of nodes. We use the variance of the residual energy of nodes to evaluate the load balance of a protocol. The more the variance, the closer the energy value of each node is, the better the load balance of the network is
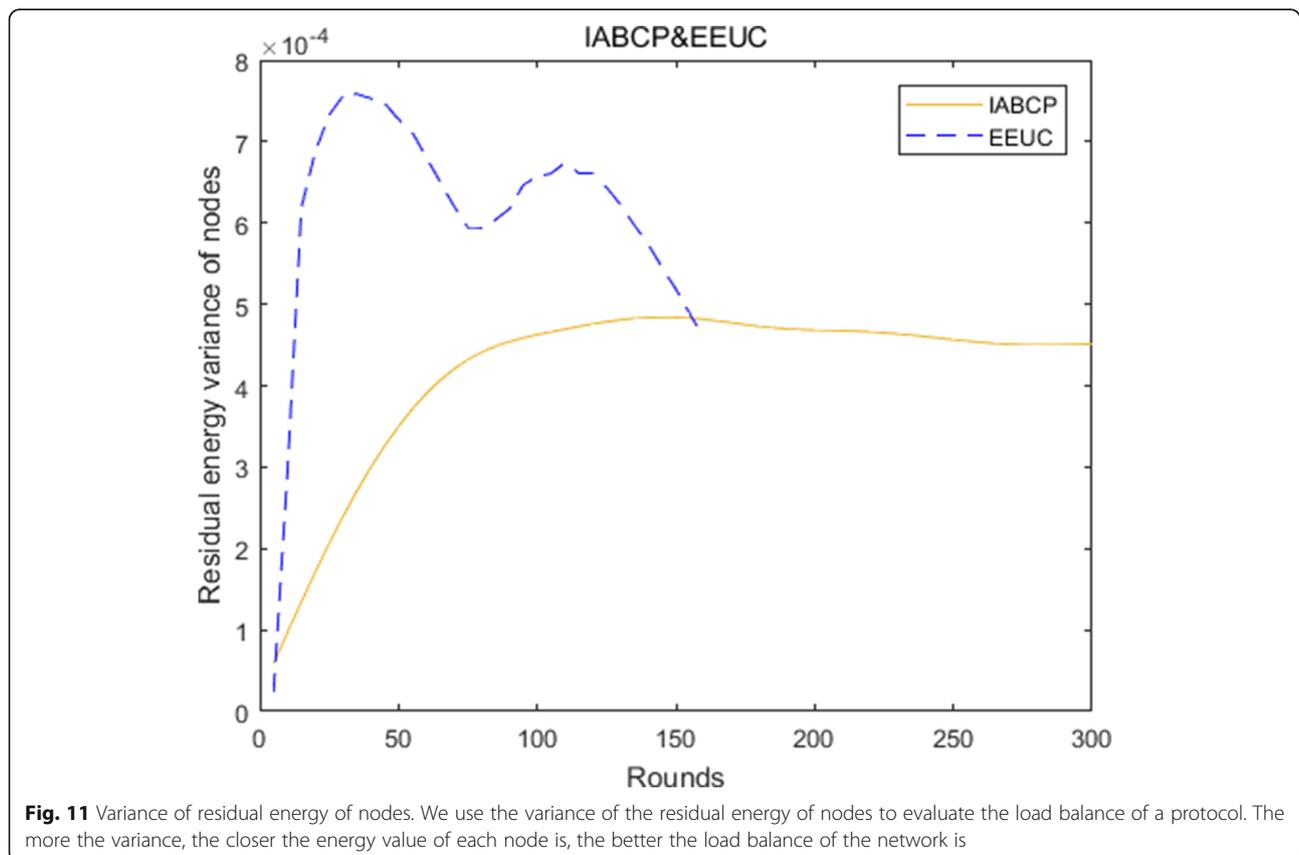
solutions with shorter routing path, larger average energy, and larger minimum energy of passing nodes have higher fitness value and have more chances to become the final choice. By this method, the solution fitness of a shorter path is higher when the energy is similar. When the energy of nodes is differentiated, the solution with higher average energy and higher minimum energy and shorter path length is preferred.

5 Simulation analysis

In order to verify the performance of this protocol, the simulation is carried out using MATLAB 2017a. The simulation conditions are shown in Table 1. In the simulation experiment, there may be changes in the number of nodes. Other conditions are completely in accordance with the simulation parameters in Table 1. The simulation contents include routing path schematic diagram of protocol construction, network life cycle comparison, network load balance comparison, and bee colony algorithm optimization analysis.

   The initial location of the sink node is located in the wireless sensor network, and the location is random. The node collects data periodically and sends the data to the sink node for communication. The moving speed of the sink node is 5 *m/s*. The cluster radius is $d_0/2$, which

makes the communication distance L between CH, $d_0/2 < L < d_0$. If the cluster radius continues to increase, the communication distance between CH and CH may exceed d_0, resulting in a rapid increase in energy consumption. In this simulation, the first node's dead time and the variance of residual energy are used to measure the load balance and average communication energy consumption of the protocol, and the lifetime of the network is measured by half of the node's dead time. In networks with similar average communication energy consumption, the first dead node with good load balance appears later, while the first dead node will appear earlier if the average communication energy consumption of the protocol itself is large or the load balance is poor. The variance coefficient can reflect the discrete degree of data, so it can be used to measure the balance of network load (discrete degree of node energy). When half of the nodes in wireless sensor networks die, the network is often unable to complete the tasks set, so it can be considered that the network has failed.

   The rotation period of the cluster head will also affect the performance of the protocol. When the rotation period of the cluster head is too long, it will damage the load balance of the network. If the rotation period of the cluster head is too short, it will increase the energy
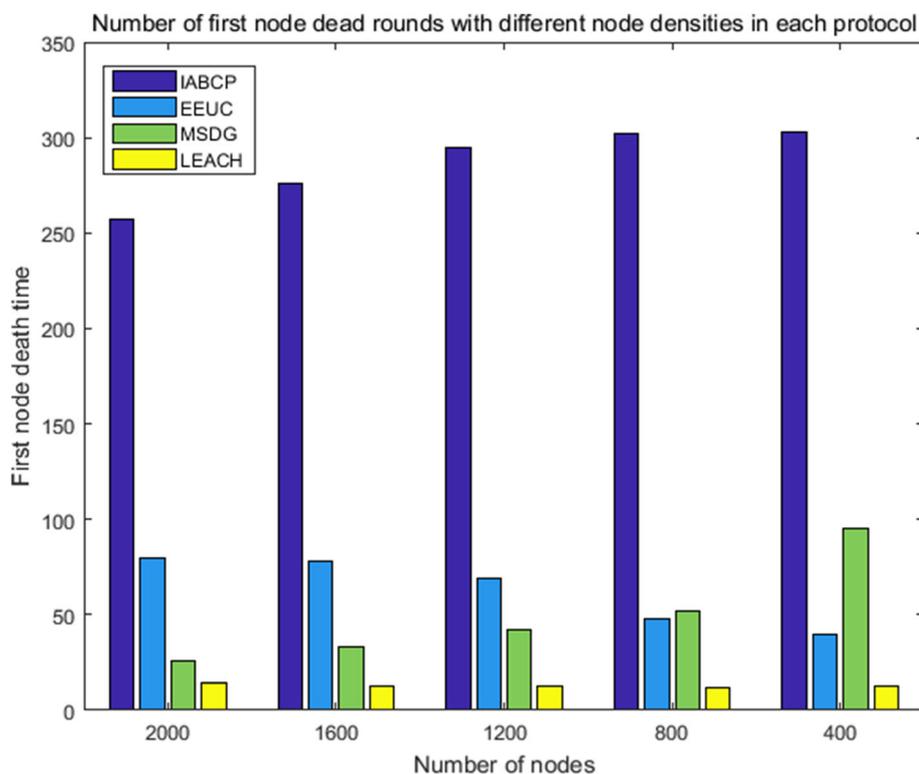

Fig. 12 Number of first node dead rounds with different node densities in each protocol. In order to prove the wide applicability of this protocol, the simulation of the number of the first dead node turns is carried out in the equal area under different node density

consumption of the network. Therefore, the protocol is simulated under different rotation periods of cluster head. The results are shown in Fig. 6.

It can be seen that when the cluster head rotation cycle is short, the first dead node appears later, but half of the nodes die earlier. When the rotation period increases, the first death node appears earlier, but half of the nodes die later. The lifetime of the network is extended, but when the rotation period is too large, the death time of the first node appears earlier, and the lifetime of the network is reduced. It can be seen from the experimental results that choosing 9–13 as the rotation period of cluster head can make the network reach a better state, so in this paper, choosing ten as the rotation period of cluster head.

5.1 Routing path
The cluster formed by the protocol and the route path constructed are shown in Fig. 7. When the sink node location does not exceed a communication radius distance of the previous update routing table location, the data is forwarded to the sink node within a hop range, and the sink node can still maintain stable communication. When the location of the sink node exceeds the

communication range with the location of the last update routing table, it is necessary for the sink node to regenerate the routing table. The cluster head receives the updated routing table and forwards the information according to the new routing table. The red line is the route constructed by the IABC algorithm. After the routing table is generated by the sink node, the routing table is forwarded to the farthest node according to the inverted method. The farthest node sends data according to the routing table. The relay node fuses its own information with the forwarding information and then forwards it according to the routing table until it is sent to the sink node.

5.2 Performance analysis of protocol
On the premise of the above simulation, the network lifetime of EEUC protocol with a mobile sink node, LEACH protocol with a mobile sink node, MSDG protocol with a mobile sink node, and our protocol is simulated. The total number of nodes is 1600. The number of surviving nodes varies with the number of rounds as shown in Fig. 8. Compared with the other three protocols, LEACH protocol with mobile sink has the shortest lifetime, while other protocols adopt multi-hop routing.
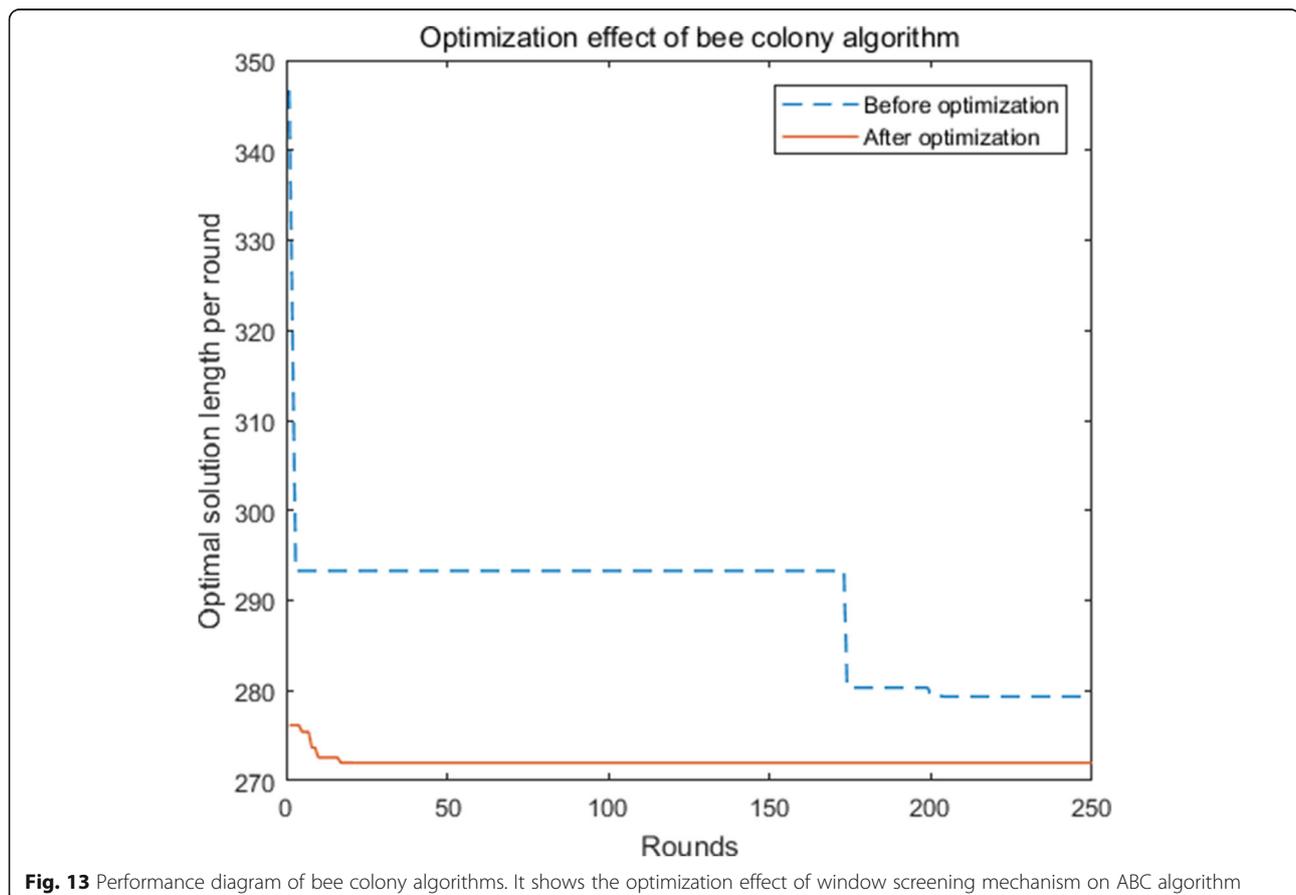

Fig. 13 Performance diagram of bee colony algorithms. It shows the optimization effect of window screening mechanism on ABC algorithm

By comparing the lifetime, we can find that it is more suitable to use multi-hop routing in large networks. Although MSDG adopts the method of multi-hop routing, in the face of high-density network environment, its load balance is poor, and its service life is short. However, the clustering method of EEUC protocol is relatively primitive, and the construction of inter-cluster communication routing needs the whole network broadcasting. The communication process is more complex, so its lifetime is shorter.

We chose one of the best performing protocols and compared its main performance with our protocol in the following papers. Figures 9, 10 and 11 show the variation curves of the number of dead nodes, the average residual energy of nodes, and the variance of residual energy of nodes for IABCP and EEUC protocols. It can be seen from Figs. 9 and 10 that half of the nodes in the IABCP protocol have run more than 400 rounds when they die, and the average residual energy of the nodes is less than 5%. Compared with the EEUC protocol, half of the nodes run 160 rounds when they die, and the average residual energy of the nodes is only about 10%. It can be concluded that the network energy consumption of our protocol is better than that of EEUC protocol. Through Fig. 11, the variance of residual energy of the IABC protocol is only one half of that of EEUC protocol, which proves that the load balance of the

IABCP protocol is much better than the EEUC protocol. Compared with EEUC, in clustering method, nodes only need to communicate with the nodes within their own communication radius, and the data transmission range of each node is small, and the energy consumption is low. And more importantly, in the inter cluster communication, this protocol adopts the method of routing generated by the sink node. Only when the routing table is generated for the first time, the location of each CH is obtained. In other cases, the sink node calculates the time when the CH-β node replaces the CH and generates the routing table and transmits it to each new CH. Therefore, the energy consumption of this protocol is very low in the process of clustering and forming the inter cluster routing path. In addition, this protocol considers the problem of network load balance through fitness function when constructing routing path, further enhances the network load balance, and reduces the residual energy variance of network nodes.

Node density also affects network lifetime of the different protocols. LEACH protocol and EEUC protocol are not significantly affected by node density, while MSDG protocol is more sensitive to node density because of its tree-forming method. In order to compare the performance changes of this protocol under various node densities, the EEUC protocol, MSDG protocol, LEACH
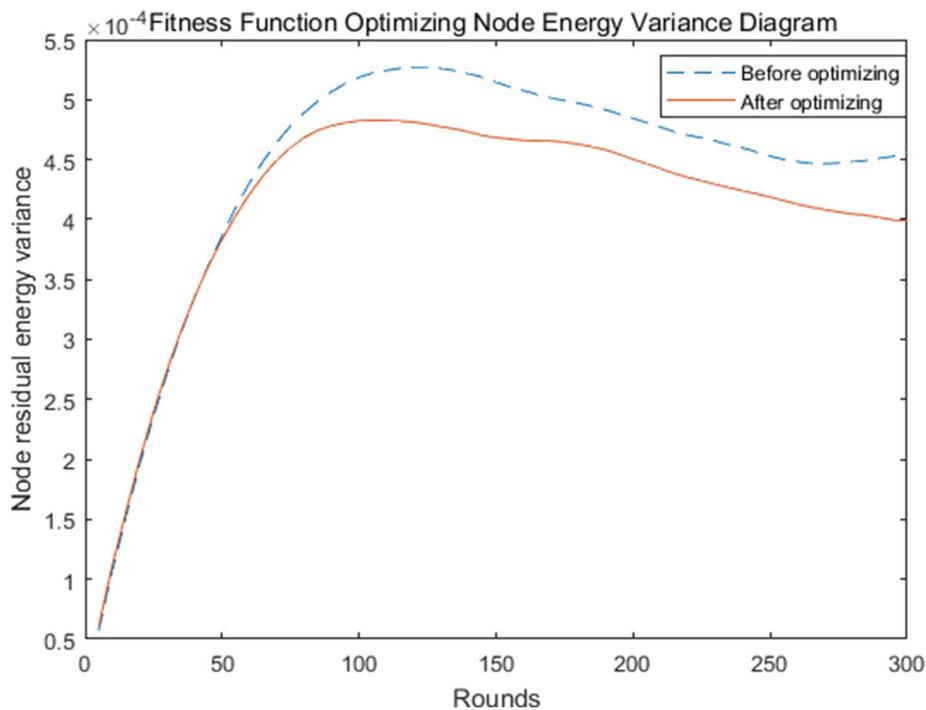

Fig. 14 Fitness function optimizing node energy variance diagram. It is proved that the improvement of fitness function optimizes the network load balance
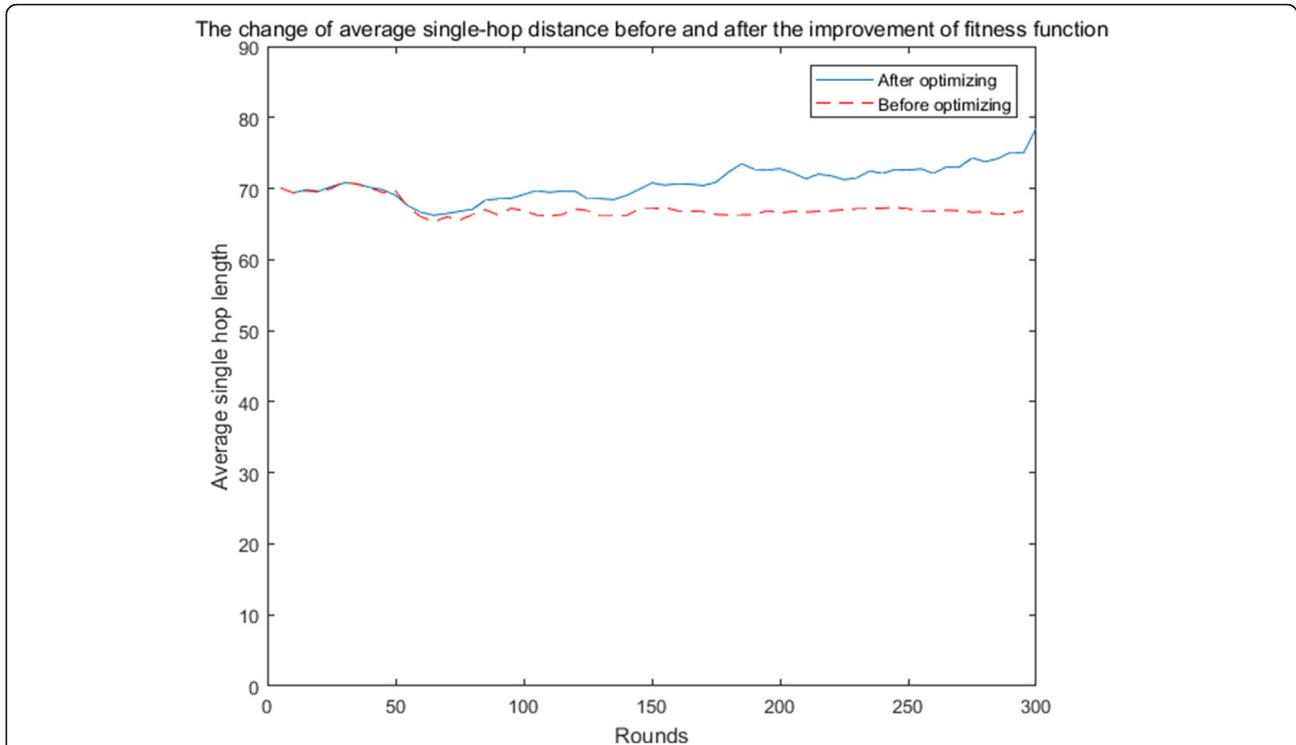
Fig. 15 The change of average hops before and after the improvement of fitness function. It shows that the improvement of fitness function affects the average single-hop distance of communication
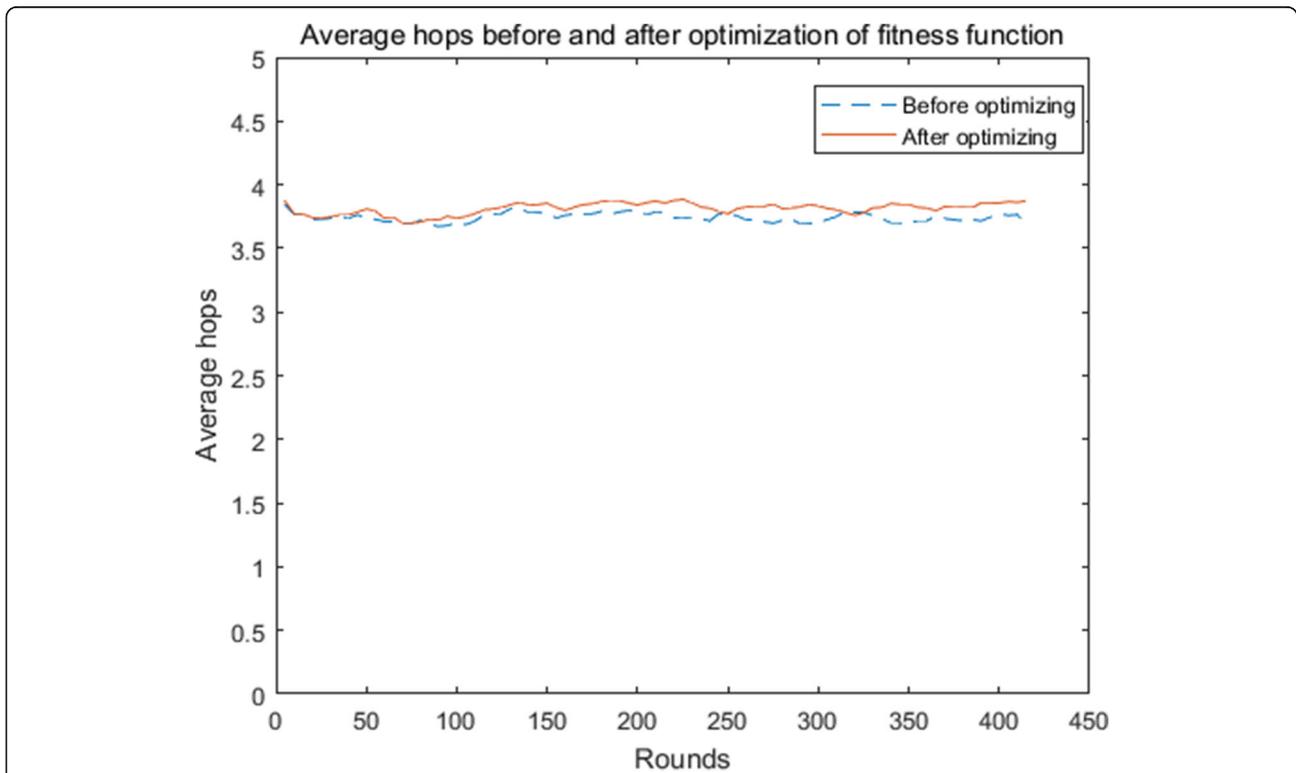

Fig. 16 The change of average single-hop distance before and after the improvement of fitness function. It shows the influence of the improvement of fitness function on the average number of communication hops

protocol, and IABCP protocol are simulated under different node densities in the same area. The first node's death time is shown in Fig. 12.

By observing the data, it can be found that the node density has less influence on the lifetime of LEACH protocol, and its lifetime is mainly affected by the size of the network layout area. The density of nodes has a significant impact on the MSDG protocol. When the density of nodes is small, the lifetime of the MSDG protocol is longer. On the contrary, the lifetime of the EEUC protocol increases with the increase of the density of nodes, because the larger the cluster head load, the more nodes can be rotated, the more nodes can alleviate the death of nodes. It can be seen from the graph that the performance of this protocol is stable when the node density changes. It basically keeps the first dead node between 250 rounds and 300 rounds. It can be seen that the adaptability of this protocol to different environments is higher than that of EEUC protocol and MSDG protocol. This protocol is superior to the other three protocols in terms of network lifetime.

5.3 Analysis of optimization performance of bee colony algorithm

If the network layering mechanism is canceled, the bee colony algorithm becomes a completely random construction solution, which is inefficient. It is obvious to compare the results of bee colony algorithm with or without layered routing mechanism. Therefore, the original bee colony algorithm is compared with the bee colony algorithm optimized by adding window screening mechanism under the same initial node. The results are shown in Fig. 13. The dotted line is the original bee colony algorithm, and the real line is the bee colony algorithm optimized by adding window screening mechanism. It can be clearly seen that the quality of the initial solution generated by the optimized bee colony algorithm and the optimization speed of the solution are greatly improved. The initial solution of the improved bee colony algorithm through the window screening mechanism is better than the result of 250 rounds of calculation before optimization. The optimization speed of the improved algorithm is much faster than that of the original algorithm. It can be proved that adding window screening mechanism in bee colony algorithm is very effective in solving the problem of constructing routing path.

Figure 14 shows the optimization effect of the fitness function. The two curves in Fig. 14 are the variation of residual energy variance of 300 rounds of 1600 nodes before and after optimization. It can be seen that the load balancing degree of the first 50 rounds is similar,
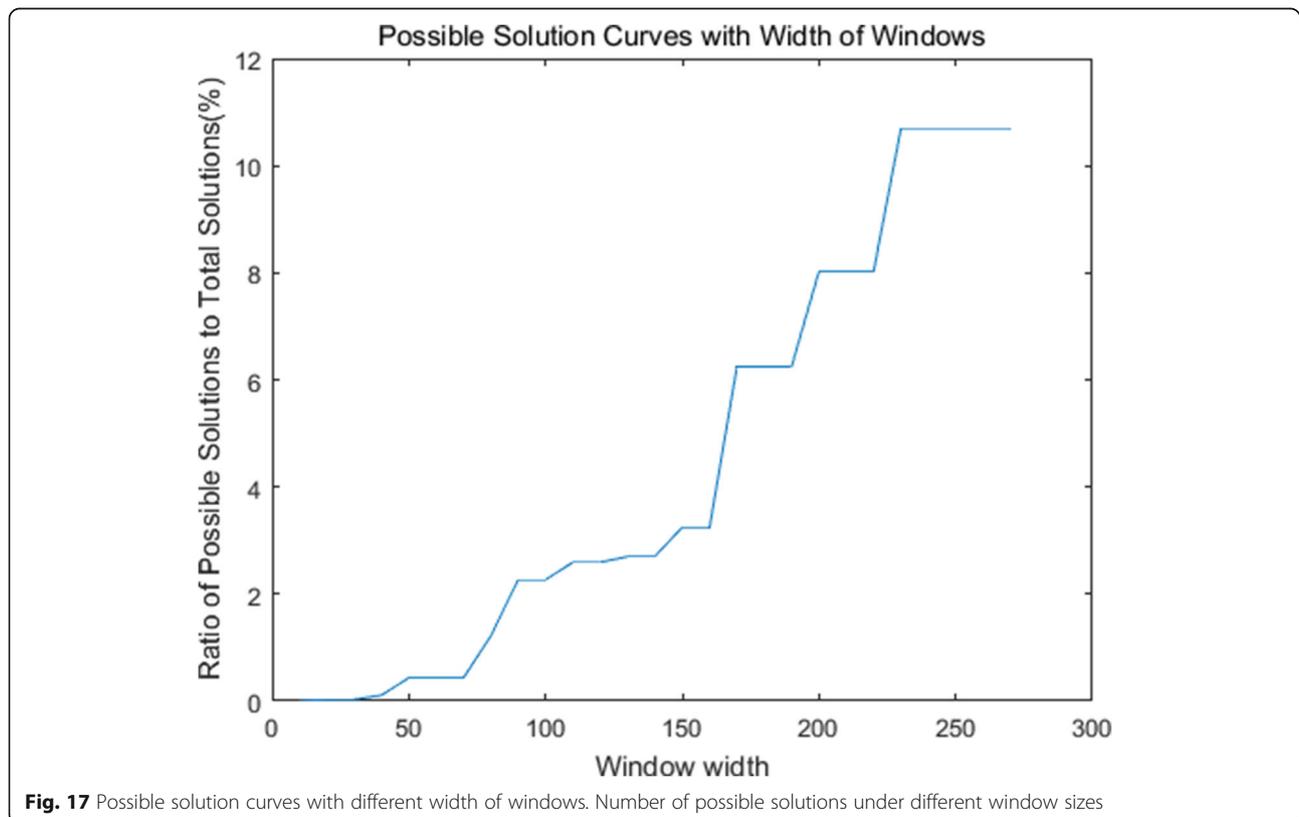

Fig. 17 Possible solution curves with different width of windows. Number of possible solutions under different window sizes

and the path length of the generated solution is the main consideration. With the increase of network operation rounds, the optimized fitness function begins to consider the cluster head energy after energy differentiation and makes the routes with higher energy have higher fitness. Therefore, compared with the fitness function before optimization, the optimized fitness function reduces the residual energy variance of nodes after 50 rounds and enhances the load balance of the network. Figures 15 and 16 are the curves of the average length and average hops of the path generated by the bee colony algorithm before and after the improvement as a function of the number of rounds. As can be seen from Fig. 15, the average length of single hop generated by the optimized bee colony algorithm increases with the increase of rounds. Because the energy of cluster head nodes differentiates with the increase of rounds, and the load balance in the fitness function becomes the main factor, the path length becomes the secondary factor. And in Fig. 16, the path length before the optimization is relatively

stable. The number of hops generated before and after optimization is basically the same and about four. It also shows that the method of generating routing in the order of cluster head distance from the sink node from far to near can effectively reduce the routing construction starting from close cluster head.

The size of windows has a great influence on the routing solution construction of bee colony algorithm. When the windows are too small, it will lead to the problem that the path cannot be formed or the single-hop distance is too large. When the windows are too large, the effect of reducing the possible solution will be reduced, and the quality of the output solution will be poor under the optimization of fixed rounds. The ratio of window size to the number of possible solutions to the total solution is shown in Fig. 17. It is obvious that the number of possible solutions varies greatly under the window sizes of 60, 90, 140, 180, 220, and 270. Therefore, the protocol is simulated under the above circumstances. The result is shown in Fig. 18, the residual energy variance of nodes with different window sizes can be seen
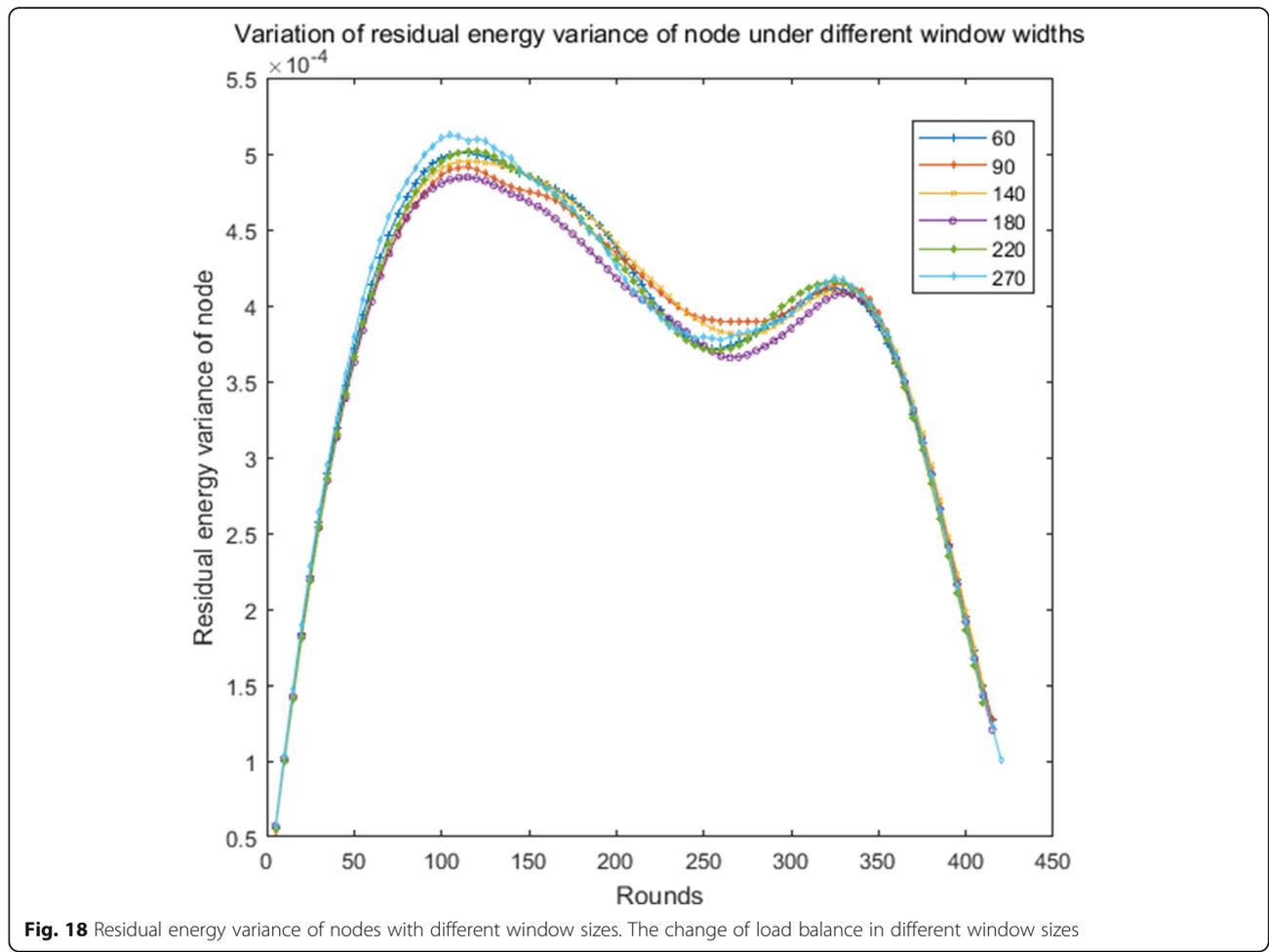

Fig. 18 Residual energy variance of nodes with different window sizes. The change of load balance in different window sizes

that the overall network load balance is the best when the window size is 180.

6 Results and discussion

A data acquisition protocol IABCP for MSWSNs is proposed in this paper. In this protocol, we propose a new routing table generation method, which gives the task of routing table generation to sink node. A routing path is generated by the sink node through IABC, which considers the energy consumption and load balance of the network. With the addition of a new sub cluster head CH-β node, this routing path construction method only needs to know the location of each cluster head when the network is clustered for the first time. By calculating the number of rounds of CH-β node replacing the cluster head, the sink node directly calculates the routing table of data transmission of each new cluster head. The energy consumption of the whole route construction process is very low. And our protocol adopts a new clustering method. Nodes determine the time of claiming cluster head by the average energy of nodes within the communication radius and their own residual energy, which reduces the energy consumption of each node in the clustering stage. The simulation results show that this protocol performs much better than LEACH, EEUC, and MSDG in communication energy consumption, network lifetime, and network load balance respectively.

Because the method of enhancing the network load balance adopted in this protocol does not conflict with the current methods such as non-uniform clustering [27, 28], we can try to combine with the non-uniform clustering in the future work to further enhance the network load balance. And the routing table generation method proposed in this paper has great development potential. It can optimize its efficiency and routing generation method in the follow-up work. And in the follow-up work, the function of CH-β node can also be added. For example, CB-β can share a part of inter cluster communication or intra cluster communication load when the communication load between cluster heads is high. The algorithm proposed in this paper is suitable for low-power ad-hoc wireless sensor networks to implement cluster head selection and routing path planning. In future work, we will apply it to actual network environments.

Abbreviations
WSNs: Wireless sensor networks; LEACH: Low-energy adaptive cluster hierarchical; EECS: Energy efficient clustering scheme; MSPR: Mobile sink-based routing protocol; CH: Cluster head; DM: Data massage; MSDG: Mobile sink-based data gathering; EEUC: Economic-environmental unit-commitment; MSWSNs: Mobile sink wireless sensor networks; TTDD: Two-tier data dissemination; HEED: Hybrid energy efficient distributed; ABC: Artificial bee colony algorithm; CM: Control massage; CH-β: Cluster head-β; RUHEED: Rotated unequal hybrid energy efficient distributed; PEGSIS: Power efficient gathering in sensor information systems

Authors' contributions
Conceived and designed the experiments: TYZ and GC. Performed the experiments: QTZ and TYZ. Analyzed the data: HD, TYZ, and CL. Wrote the paper: TYZ, GC, and QTZ. All authors have contributed to this research work and read and approved the final manuscript.

Authors information
Tianyi Zhang received the M.S. degree in Measurement and Control Technology and Instruments from Shijiazhuang Tiedao University, Shijiazhuang, China, in 2017. His current research interests include Wireless Sensor Network and Deep Net learning.
Geng Chen received the B.S. degree in electronic information engineering and the M.S. degree in communication and information system from Shandong University of Science and Technology, Qingdao, China, in 2007 and 2010, respectively, and the Ph.D. degree in information and communications engineering from Southeast University, Nanjing, China, in 2015. He is currently an Associate Professor with the College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao, China. His current research interests are in the areas of heterogeneous networks, ubiquitous networks, and software defined mobile networks, with emphasis on wireless resource management and optimization algorithms, and precoding algorithms in large scale MIMO.
Qingtian Zeng received the B.S. degree and the M.S. degree in computer science from Shandong University of Science and Technology, Taian, China, in 1998 and 2001 respectively, and the Ph.D. degree in computer software and theory from the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China, in 2005. He is currently a Professor with Shandong University of Science and Technology, Qingdao, China. His research interests are in the areas of Petri nets, process mining, and knowledge management.
Ge Song received the master's degree in computer science and technology from Shandong University of Science and Technology, Shandong, China in 2004. He is currently teaching in the College of Electronic Information Engineering. His current research interests include Internet of Things, Wireless Sensor Networks, Robot Path Planning, and Deep Learning.
Chao Li received the B.S. and M.S. degrees from the Shandong University of Science and Technology and the Ph.D. degree from the Chinese Academy of Sciences, in 2014. He was a Visiting Scholar with The Hong Kong University of Science and Technology, from 2014 to 2015. He is currently a Lecturer with the Shandong University of Science and Technology. His research interests include social media, natural language processing, data mining, and network embedding learning. He is a member of CCF.
Hua Duan received the B.S. and M.S. degrees in applied mathematics from the Shandong University of Science and Technology, Tai'an, China, in 1999 and 2002, and the Ph.D. degree in applied mathematics from Shanghai Jiao Tong University, in 2008. She is currently a Professor with the Shandong University of Science and Technology. Her research interests include process mining and machine learning.

Availability of data and materials
Not applicable.

Author details
[1]College of Electronic and Information Engineering, Shandong University of Science and Technology, Qingdao 266590, China. [2]College of Mathematics and System Science, Shandong University of Science and Technology, Qingdao 266590, China.

References
1. W.R. Heinzelman, A. Chandrakasan, H. Balakrishnan, *Energy-efficient communication protocol for wireless microsensor networks, 10* (2000)
2. Wang, N., Zhu, H.: An energy efficient algrithm based on leach protocol. In: 2012 International Conference on Computer Science and Electronics Engineering, vol. 2, pp. 339–342 (2012).
3. Z. Cui, Y. Cao, X. Cai, J. Cai, J. Chen, Optimal leach protocol with modified bat algorithm for big data sensing systems in internet of things. J Parallel Distributed Comput **132**, 217–229 (2019)
4. L. Zhang, K. Zhang, J. Song, Mobile sink-based data gathering algorithm in wireless sensor networks. Chin J Sensors Actuators **5**, 024 (2012)
5. Y. Pal Verma, A. Kumar, Economic-emission unit commitment solution for wind integrated hybrid system. Int J Energy Sector Manag **5**(2), 287–305 (2011)
6. J. Wang, J. Cao, R.S. Sherratt, J.H. Park, An improved ant colony optimization-based approach with mobile sink for wireless sensor networks. J Supercomputing **74**(12), 6633–6645 (2018)
7. Wang, T., Zeng, J., Lai, Y., Cai, Y., Tian, H., Chen, Y., Wang, B.: Data collection from wsns to the cloud based on mobile fog elements. Future Generation Computer Systems (2017)
8. J.-S. Pan, L. Kong, T.-W. Sung, P.-W. Tsai, V. Snáˇsel, A clustering scheme for wireless sensor networks based on genetic algorithm and dominating set. J Internet Technol **19**(4), 1111–1118 (2018)
9. Tokala, M., Nallamekala, R.: Secured algorithm for routing the military field data using dynamic sink: Wsn. In:2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT),pp. 471–476 (2018).
10. F. Al-Turjman, Cognitive routing protocol for disaster-inspired internet of things. Future Generation Computer Systems **92**, 1103–1115 (2019)
11. F.A. Khan, M. Khan, M. Asif, A. Khalid, I.U. Haq, Hybrid and multi-hop advanced zonal-stable election protocol for wireless sensor networks. IEEE Access **7**, 25334–25346 (2019)
12. S. Din, A. Paul, A. Ahmad, J.H. Kim, Energy efficient topology management scheme based on clustering technique for software defined wireless sensor network. Peer-to-Peer Networking and Applications **12**(2), 348–356 (2019)
13. Y. Xue, J. Jiang, B. Zhao, T. Ma, A self-adaptive artificial bee colony algorithm based on global best for global optimization. Soft Computing **22**(9), 2935–2952 (2018)
14. A. Al-Baz, A. El-Sayed, A new algorithm for cluster head selection in leach protocol for wireless sensornetworks. Int J Commun Syst **31**(1), 3407 (2018)
15. K.M. Modieginyane, B.B. Letswamotse, R. Malekian, A.M. Abu-Mahfouz, Software defined wireless sensor networks application opportunities for efficient network management: A survey. Comput Electrical Eng **66**, 274–287 (2018)
16. S.K. Singh, P. Kumar, J.P. Singh, A survey on successors of leach protocol. IEEE Access **5**, 4298–4328 (2017)
17. S. Chand, S. Singh, B. Kumar, Heterogeneous heed protocol for wireless sensor networks. Wireless personal communications **77**(3), 2117–2139 (2014)
18. Ullah, Z., Mostarda, L., Gagliardi, R., Cacciagrano, D., Corradini, F.: A comparison of heed based clustering algorithms–introducing er-heed. In: 2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA), pp. 339–345 (2016).
19. Wang, J., Gao, Y., Yin, X., Li, F., Kim, H.-J.: An enhanced pegasis algorithm with mobile sink support for wireless sensor networks. Wireless Communications and Mobile Computing 2018 (2018)
20. Nazir, B., Hasbullah, H.: Mobile sink based routing protocol (msrp) for prolonging network lifetime in clustered wireless sensor network. In: 2010 International Conference on Computer Applications and Industrial Electronics,pp. 624–629 (2010).
21. C. Wang, S. Wang, Research on Uneven Clustering APTEEN in CWSN Based on Ant Colony Algorithm [J]. IEEE Access (2019)
22. A. Somauroo, V. Bassoo, Energy-efficient genetic algorithm variants of PEGASIS for 3D Wireless Sensor Networks [J]. Appl Comput Info (2019)
23. M. Wazid, A.K. Das, S. Kumari, M.K. Khan, Design of sinkhole node detection mechanism for hierarchical wireless sensor networks. Security and Communication Networks **9**(17), 4596–4614 (2016)
24. D. Wu, J. He, H. Wang, C. Wang, R. Wang, A hierarchical packet forwarding mechanism for energy harvesting wireless sensor networks. IEEE Communications Magazine **53**(8), 92–98 (2015)
25. S. Song, W. Zhang, P. Han, D. Zou, Sliding window method for vehicles moving on a long track. Vehicle system dynamics **56**(1), 113–127 (2018)
26. A.N.Z. Rashed, A. Mohammed, H.A. Sharshar, A.M. El-Eraki, Fast routing algorithm in optical multistage interconnection networks using fast window method. Int J Advanced Res Electron Commun Eng **6**(1), 37–43 (2017)
27. J. Kasza, K. Hemming, R. Hooper, J. Matthews, A. Forbes, Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials. Stat Methods Med Res **28**(3), 703–716 (2019)
28. Hanasaki, I., Hosokawa, C.: Non-uniform stochastic dynamics of nanoparticle clusters at a solid–liquid interface induced by laser trapping. Japanese Journal of Applied Physics 58(SD), 07 (2019)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.